

tubes

Kelompok 1 RA
2025

Penerapan Rantai Markov untuk Menentukan Probabilitas Kategori Hujan Harian di ITERA

```
library(readr) # Untuk membaca CSV
library(dplyr) # Untuk manipulasi data (filter)

## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##   filter, lag

## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union

library(lubridate) # Untuk bekerja dengan tanggal (memfilter tahun)

## Warning: package 'lubridate' was built under R version 4.2.3

## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##   date, intersect, setdiff, union

library(markovchain) # Untuk analisis Rantai Markov

## Warning: package 'markovchain' was built under R version 4.2.3

## Package: markovchain
## Version: 0.9.5
## Date: 2023-09-24 09:20:02 UTC
## BugReport: https://github.com/spedygiorgio/markovchain/issues

## Attaching package: 'markovchain'

## The following object is masked from 'package:lubridate':
##   period

library(knitr)
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.2.3

library(forcats)

data <- "D:\College\9th\PenStok\tubes\data_harian_terklasifikasi.csv"

# Membaca data lengkap
data_full <- read_csv(data, show_col_types = FALSE)

# Filter tahun 2024
data_full %>% mutate(Datetime = ymd(Datetime)) %>% # Konversi kolom ke format Tanggal
filter(year(Datetime) == 2024) # Ambil data tahun 2024

if (nrow(data_2024) < 2) {
  cat("Error: Data 2024 tidak ditemukan atau tidak cukup untuk analisis.")
} else {
  cat(paste("Data 2024 berhasil difilter. Jumlah observasi:", nrow(data_2024), "\n\n"))

  # Ambil urutan keadaan (state) dari data 2024
  state_sequence_2024 <- data_2024$state

  cat("Contoh data 2024 yang digunakan:")
  kable(head(data_2024), caption = "Data Cuaca Tahun 2024 (6 Baris Pertama)")
}

## Data 2024 berhasil difilter. Jumlah observasi: 244
##
## Contoh data 2024 yang digunakan:
```

Data Cuaca Tahun 2024 (6 Baris Pertama)

Datetime total_rainfall_harianstate

2024-01-01 7,6Hujan Ringan

2024-01-02 17,4Hujan Ringan

2024-01-03 411,4Hujan Sangat Lebat

2024-01-05 3308,2Hujan Sangat Lebat

2024-01-06 1150,0Hujan Sangat Lebat

12,8Hujan Ringan

cat("Ringkasan Statistik untuk Total Curah Hujan Harian (mm/hari) 2024:\n")

Ringkasan Statistik untuk Total Curah Hujan Harian (mm/hari) 2024:

ringkasan_numerik <- data_2024 %>%
 select(total_rainfall_harian) %>%
 summarize()

kable(ringkasan_numerik, caption = "Statistik Deskriptif Curah Hujan Harian 2024")

Statistik Deskriptif

Curah Hujan Harian

2024

total_rainfall_harian

Min.: 0.0

1st Qu.: 0.0

Median: 20.6

Mean: 345.6

3rd Qu.: 272.4

Max: 5812.4

cat("Tabel Frekuensi Kategori Cuaca (State) 2024:\n")

Tabel Frekuensi Kategori Cuaca (State) 2024:

Hitung frekuensi dan persentase
urutan_state_logis <- c("Berawan", "Hujan Ringan", "Hujan Sedang", "Hujan Lebat", "Hujan Sangat Lebat")
urutan_state_logis

for(i in 1:length(urutan_state_logis)) {
 tabel_frekuenси <- data_2024 %>%
 count(state, name = "Jumlah Hari") %>%
 mutate(Persentase = round((Jumlah_Hari / sum(Jumlah_Hari)) * 100, 2)) %>%
 # Mengatur urutan state sesuai urutan logis
 mutate(state = factor(state, levels = urutan_state_logis)) %>%
 arrange(state)

Tampilkan tabel menggunakan kable
kable(tabel_frekuenси,
 caption = "Tabel Frekuensi Kategori Cuaca Mariano 2024",
 col.names = c("Kategori (State)", "Jumlah Hari", "Persentase (%)"))

Tabel Frekuensi Kategori Cuaca Harian 2024

Kategori (State) Jumlah Hari Persentase (%)

Berawan 102 41.80

Hujan Ringan 19 7.79

Hujan Sedang 29 11.69

Hujan Lebat 9 3.69

Hujan Sangat Lebat 85 34.84

Atur tampilan default untuk semua plot
theme_set(theme_minimal())

(Bar Chart) Untuk Frekuensi Kategori

plot_bar <- ggplot(tabel_frekuenси, aes(x = state, y = Jumlah_Hari, fill = state)) +
 geom_bar(stat = "identity", show.legend = FALSE) +
 geom_text(aes(label = paste0(Jumlah_Hari, "%")))

labs(title = "Distribusi Kategori Cuaca (State) 2024",
 x = "Kategori (State)",
 y = "Jumlah Hari (Frekuensi)") +
 theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Putar label jika panjang

print(plot_bar)

Distribusi Kategori Cuaca (State) 2024

Jumlah Hari (Frekuensi)

100

85 hari

60 hari

40 hari

20 hari

10 hari

0

Berawan Hujan Ringan Hujan Sedang Hujan Lebat Hujan Sangat Lebat

Kategori (State)

Grafik Runut Waktu (Time Series Plot)

plot_ts <- ggplot(data_2024, aes(x = Datetime, y = total_rainfall_harian)) +
 geom_line(color = "navy", alpha = 0.7) +
 # Tandai hari dengan curah hujan lebat (misal > 50 mm/hari)
 geom_point(data = filter(data_2024, total_rainfall_harian > 50),
 aes(x = Datetime, y = total_rainfall_harian),
 color = "red", size = 2)

labs(title = "Curah Hujan Harian Sepanjang Tahun 2024",
 subtitle = "Titik merah menandakan hujan lebat (> 50 mm/hari)",
 x = "Tanggal",
 y = "Total Curah Hujan (mm/hari)")

print(plot_ts)

Curah Hujan Harian Sepanjang Tahun 2024

Titik merah menandakan hujan lebat (> 50 mm/hari)

6000

4000

2000

0

Jan Apr Jul

Tanggal

2. Identifikasi Ruang Keadaan (State Space)

Ruang keadaan adalah himpunan semua status cuaca unik yang terjadi pada tahun 2024.

```
unique_states <- unique(state_sequence_2024)
cat("Ruang Keadaan (State Space) 2024:\n")
```

```
## Ruang Keadaan (State Space) 2024:
```

```
print(unique_states)
```

```
## [1] "Hujan Ringan" "Hujan Sangat Lebat" "Hujan Sedang"
## [4] "Hujan Lebat" "Berawan"
```

3. Matriks Transisi

Kita mengestimasi matriks transisi 1-langkah (P^n) dari data. P_{ij} adalah probabilitas berpindah dari keadaan i ke keadaan j pada hari berikutnya.

```
# Mengestimasi matriks transisi dari data 2024
mc_fit_2024 <- markovchainFit(data = state_sequence_2024)
```

```
# Matriks probabilitas transisi  $P_{ij}(t)$ 
transition_matrix_2024 <- mc_fit_2024$estimate@transitionMatrix
```

```
cat("Matriks Transisi (P) 2024:\n")
```

```
## Matriks Transisi (P) 2024:
```

```
kable(round(transition_matrix_2024, 3), caption = "Matriks Transisi Probabilitas (P) 1-Langkah")
```

Matriks Transisi Probabilitas 1-Langkah

	Berawan	Hujan Lebat	Hujan Ringan	Hujan Sangat Lebat	Hujan Sedang
Berawan	0.467	0.030	0.030	0.281	0.059
Hujan Lebat	0.373	0.111	0.000	0.000	0.444
Hujan Ringan	0.474	0.000	0.053	0.368	0.105
Hujan Sangat Lebat	0.165	0.035	0.141	0.482	0.176
Hujan Sedang	0.276	0.069	0.034	0.552	0.069

```
# Membuat objek markovchain
mc_cuaca_2024 <- new("markovchain",
  states = unique_states,
  transitionMatrix = transition_matrix_2024,
  name = "Model Cuaca 2024")
```

```
cat("nRingkasan Model Rantai Markov:\n")
```

```
## Ringkasan Model Rantai Markov:
```

```
summary(mc_cuaca_2024)
```

Model Cuaca 2024 Markov chain that is composed by:

Closed classes:

Hujan Ringan,Hujan Sangat Lebat,Hujan Sedang,Hujan Lebat,Berawan

Recurrent classes:

(None)

The Markov chain is Irreducible

The absorbing states are: NONE

Kesimpulan: Rantai ini adalah Irreducible (Irreducible & Aperiodic). Artinya, Distribusi Stasioner valid dan unik.

```
# Cek state merayap (Aperiodic)
is_irreducible <- is_irreducible(mc_cuaca_2024)
```

```
cat(paste("n1. Apakah Rantai ini adalah Irreducible? (Semua state terhubung):", is_irreducible))
```

1. Apakah Rantai Irreducible? (Semua state terhubung): TRUE

```
# Cek Periodi (I - Aperiodic)
period_mc <- period(mc_cuaca_2024)
```

```
cat(paste("n2. Periode Rantai (jika 1, berarti Aperiodic);", period_mc))
```

Kesimpulan: Rantai ini adalah Irreducible (Irreducible & Aperiodic). Artinya, Distribusi Stasioner valid dan unik.

```
# Cek state merayap (Transient)
is_transient <- is_transient(mc_cuaca_2024)
```

```
cat(paste("n3. State Merayap (Transient States):\n"))
```

3. State Merayap (Transient States):

```
print(as.data.frame(mc_cuaca_2024))
```

character(0) ## Analisis:

• State Merayap (Transient States): character(0)

Artinya: Model Anda tidak memiliki state merayap. State merayap adalah state yang jika sekali dimasuki, tidak akan pernah kembali lagi.

• State Sementara (Transient States): character(0)

Artinya: Model Anda tidak memiliki state sementara. State sementara adalah state yang pada akhirnya akan tinggal dan tidak akan pernah kembali lagi.

7. Simulasi

7.1 Probabilitas Langkah ke-n (P^n)

Ini adalah simulasi probabilitas berpindah dari keadaan i ke j dalam n langkah (hari). Mari kita hitung untuk $n = 3$.

```
# Menghitung probabilitas 3-langkah ( $P^n$ )
n_langkah <- 3
mc_3_langkah <- mc_cuaca_2024 ^ n_langkah
```

```
# Mengambil matriks
p_3_langkah <- mc_3_langkah@transitionMatrix
```

```
cat(paste("n4. Matriks Transisi 3-langkah:\n"))
```

```
## Matriks Transisi 3-langkah:
caption = paste("Matriks Transisi Probabilitas", n_langkah, "-Langkah")
```

Matriks Transisi Probabilitas 3-langkah (P^n)

	Berawan	Hujan Lebat	Hujan Ringan	Hujan Sangat Lebat	Hujan Sedang
Berawan	0.467	0.030	0.030	0.281	0.059
Hujan Lebat	0.373	0.111	0.000	0.000	0.444
Hujan Ringan	0.474	0.000	0.053	0.368	0.105
Hujan Sangat Lebat	0.165	0.035	0.141	0.482	0.176
Hujan Sedang	0.276	0.069	0.034	0.552	0.069

```
# Mengambil probabilitas spesifik
tryCatch({
  prob_mc <- p_3_langkah["Berawan", "Hujan
```